Audio Engineering Society

Convention e-Brief

Presented at the 136th Convention
2014 April 26–29 Berlin, Germany

This Engineering Brief was selected on the basis of a submitted synopsis. The author is solely responsible for its presentation, and the AES takes no responsibility for the contents. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Audio Engineering Society.

Energy based traffic density estimation using embedded audio processing unit

György Nagy, Rene Rodigast, and Danilo Hollosi

Fraunhofer Institute for Digital Media Technology, Ilmenau, Ehrenbergstrasse 31, 98693, Germany
nagy@idmt.fraunhofer.de

ABSTRACT

In this paper, we present preliminary results of an audio-based traffic density estimation application, developed within the EU-FP7 project EAR-IT [1]. The algorithm exploits that the energy of environmental noise, generated by vehicles, is related to the prevalent traffic conditions. Noise analysis and derived restrictions were made to improve the solution, which was implemented on an embedded platform. This approach follows the current trends - distributed and local processing - and directly targets the requirements for smart cities and wireless sensor networks. Using traffic monitoring wireless sensors, provided by the testbed SmartSantander [2], development setup was established to support the audio related algorithm deployment, testing and assessment.

1. INTRODUCTION

1.1. Background

In the past few years the number of smart buildings and smart cities increased considerably [10]. Wireless Sensor Networks (WSN) in smart cities contribute to global sensing and/or situation-aware applications. The social and economic impact of such innovative applications and devices is seen to form a huge emerging market in the near future and is gaining attention all over the world. Unfortunately, acoustic sensors remain underrepresented in this context, although audio sensing is highly beneficial to expand the already available solutions in terms of situation-awareness and global sensing [5].
Over the course of the EAR-IT project [1], an embedded platform based audio processing unit (APU) is developed, which is able to process complex algorithms with high quality audio. The goal is to deploy an advanced audio sensor with complex applications that is cheap, robust, easy to install and compatible with the underlying smart city environment.

1.2. Motivation

Due to the increased motorization and urbanization, traffic jams and congestions cause problems in urban areas. They increase the fuel consumption, air and noise pollution, the transportation become slower and it is a serious stress factor in people's life [9]. With intelligent and effective traffic control, the number and the duration of jams could be decreased [4]. There are a lot of already existing technology to get to know the actual road conditions (inductive loop, radar, video, ultrasonic, etc. [3]), but these devices are very expensive or their installation is difficult. Contrary, audio based sensors do not require line-of-sight, are more flexible in terms of

installation, allow for omnidirectional sensing and are basically independent from weather conditions and lighting situations.
Advanced audio sensors and applications on embedded platforms could reveal new opportunities in smart cities.
In this paper, we present preliminary results of an audio-based traffic density estimation application, which outline the possibilities in this technology.
In Section II., a brief overview about the used audio processing unit, the first candidate of the new generation audio sensors, which is able to process complex algorithms with high quality, is given. General embedded audio preprocessing is introduced in Section III. Section IV explains the development steps of the traffic density estimation algorithm and the evaluation is going to be highlighted in Section V. The paper ends with a concluding section and a short look on the future plans.

2. AUDIO PROCESSING UNIT

The hardware selection was based on the following criteria:

- Low cost and small sensor;
- Proper audio quality to be able to deal with complex algorithms;
- Low energy consumption;
- Compatible with existing WSN (wireless sensor network) systems.

To fulfil these requirements, general embedded platform and peripherals were chosen.

In Figure 1, you can see the development hardware setup. It contains an electrets microphone (Shure MX183) with a phantom power adaptor, a preamplifier combined with a power unit over USB (t.bone MicPlug USB) and the processing board (BeagleBoard XM).
The peripherals were chosen attentively to the price, size, quality and robustness. The microphone was already tested by the manufacture outdoor for years without any malfunction, which is required in outdoor applications. The hardware structure is modular, the peripherals are changeable. If different audio quality is needed, new component could be easily added, it requires just minor changes in the software, which guarantee the sensor flexibility.

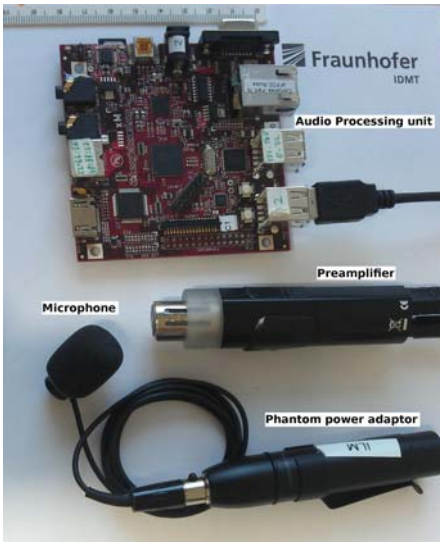


Figure 1 Development hardware setup

According to the IERC (European Research Cluster on the Internet of Things) definition of internet of thing (IoT) [8], IoT includes any smart thing able to be globally interconnected and to exchange information. The developed APU fulfills all of theses features, so the audio sensor can be considered as an IoT.
The Linux operating system, the advanced processing and the audio related performance shows the different approach of our group, which could be an innovative way of thinking about the new generation sensors. Table 1 presents the huge differences between the introduced audio sensor and the currently used IoT-s.

	APU	Waspmote [6]	Advanticsys [7]
CPU frequency	1GHz	14 MHz	8 MHz
Memory	512MB	8 KB	10 KB
Peripherals	Flexible	Fix	Fix
OS	Linux	Tiny OS	Tiny OS
Quantization	24 bits	10 bits	10 bits
Sampl. Freq.	48 kHz	8 kHz	8 kHz

Table 1 IoT, APU features

3. AUDIO PREPROCESSING ON THE EMBEDDED PLATFORM

Traffic noise analysis was done to get to know the useful audio properties and the requirements to the

developed application. Restrictions were made, regarding to noise analysis results, to decrease the computation complexity, which is crucial for power efficient, autonomous embedded platforms.

The audio sampling and preprocessing consists four major modules, namely the analog/digital conversion, DC offset removal, feature extraction and energy index calculation. The output of the preprocessing module generates the input of the estimation algorithm.

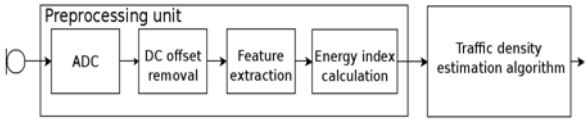


Figure 2 Audio preprocessing block diagram

The audio peripheral is attached to the line-in input of the BeagleBoard XM, which is connected to a TWL4030 module, that is an integrated power-management/audio codec device intended for use along with the ARM processor. This module makes the A/D conversion and a PCM digital audio interface is used to get the digitized data in the next software module. The sampling frequency is 44100 Hz and the quantization is 24 bits. The digitized audio is captured in 512 samples frames (~11,6ms), which is forwarded to the mean-subtraction DC removal module, then to the feature extraction, that works on Hann windowed, not overlapped frames. The next step is an energy index E_i calculation from the extracted power spectrum. The energy index is related to the loudness of the environmental noise on a restricted range; it simplifies the calculation complexity and increases the robustness.

$$E_i = \sum_k \log(X_i[k]) \qquad (1)$$

, where k is related to the bins in the spectrum, and X is the power spectrum, which is calculated from the i-th audio frame. The explained energy index is used as an input of the traffic density estimation algorithm.

4. TRAFFIC DENSITY ESTIMATION

To use a deterministic approach, the goal was to discover the relation between the environmental noise and the current traffic density, then use the noise to estimate the road saturation level. During the development process, more than two months training data was used to create a deterministic model.

4.1. Exploit global relation

To discover the relation between the number of passing cars and the energy index, data fusion, synchronization and smoothing is required. The data logging rate of the ground truth car counting IoT-s is one sample per minute, so the energy index is calculated as an arithmetic mean of the values during a minute to get the same resolution. To find an appropriate relation between the two data sets, regression analysis was done on smoothed data, which results in a high correlation and a reliable regression function. The used smoothing method is a simple moving average (SMA), which was applied on the energy index array with different window sizes n= [5, 10, 15, 20, 40, 120] samples.

$$\overline{E_i} = \frac{E_1 + E_2 + \ldots + E_n}{n}, i = [1 \ldots n], \qquad (2)$$

where n is the size of the window and E_i is the energy index related to the loudness of a minute.

To determine how strong the relation is between the predefined energy index and the current traffic density, autocorrelation coefficient was calculated on the smoothed data. Using different moving average window lengths the correlation coefficient is between 0.75 and 0.95 depends on the window size.
The high correlation proves the concept that the audio based traffic density estimation is possible.

4.2. Regression analysis

Regression analysis is a statistical process for estimating the relationships among variables, which is in our case the environmental noise and the traffic density. To visualize the connection between the data sets, the smoothed arrays were plotted. It is clearly noticeable that the relation is not linear, more second or higher order.
The plot about the data combination can be seen bellow. The x-axis is the energy index, which is related to the street noise and the y-axis it the discrete car number, which passed during a minute.

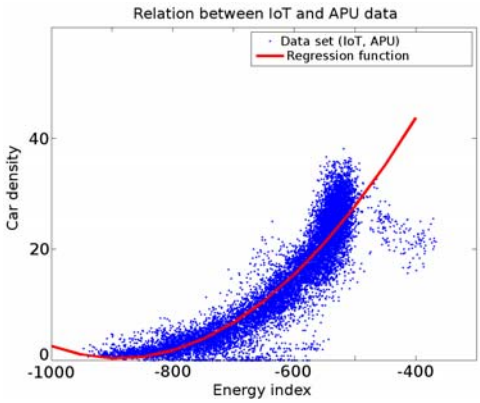


Figure 3 Relation between street noise and car density (SMA was used with 20 samples window size)

Different regression functions were examined and calculated. Finally an eighth order polynomial is used in combination with a second order one, to get the most reliable output. For choosing the best regression, 1-norm distance was used on the data sets. The lowest distanced polynomial combination was used later in the evaluation process.

In Figure 4 is shown, the already used regression function on the energy index without smoothing and with smoothing in combination with the car numbers on the street.

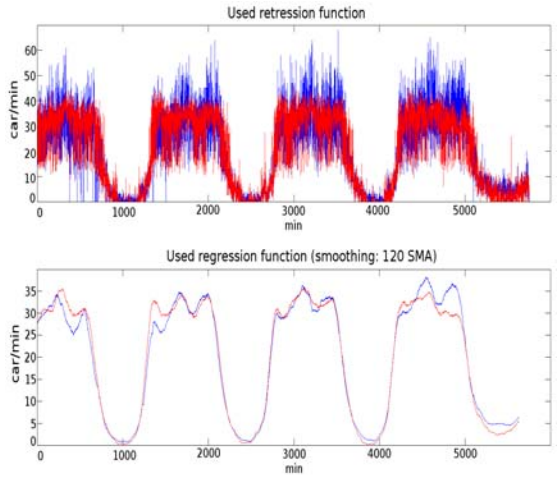


Figure 4 Used regression function and the smoothed data (blue: car numbers, red: estimated values)

At the end the output of the traffic density estimation algorithm is a level, which is related to the road saturation. The level definition is applied on the data, which is shown in Figure 4. The output range is from zero ("no cars") to five ("traffic jam").

Level	Boundary(car/min)
0	$x \leq 6$
1	$6 < x \leq 12$
2	$12 < x \leq 19$
3	$19 < x \leq 26$
4	$26 < x \leq 37$
5	$37 < x$

Table 2 Level definition

5. ALGORITHM EVALUATION

The regression function and the level definitions are applied on the energy index data set to get the estimated traffic density level. This output is used in the evaluation method, combined with the ground truth data form the car counting sensors.

5.1. Confusion matrix

The performance test is based on a confusion matrix of the actual road saturation level and the predicted audio based traffic density level. This matrix gives the opportunity to analyze the capabilities of the estimation and find out uncertainties in the algorithm. The columns represent the actual and the rows the estimated traffic density levels, which results a 6x6 matrix. The matrix elements are incremented in the position of the ground truth and the algorithm output combination.

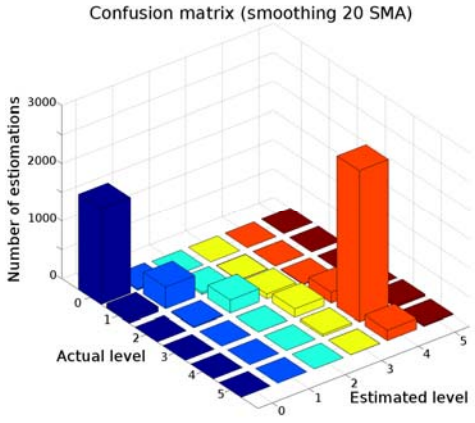


Figure 5 Confusion matrix on the smoothed data (20 moving averaged)

Confusion matrix was created following the incrementation process. Finally evaluation was done to get to know the performance of the algorithm. .

5.2. Evaluation and error calculation

Evaluation matrix was generated from four days record, which contains 5730 samples of APU and IoT data, then evaluation values were calculated form the contingency tables. The results are in Table 3.

Precision	87.83%
Sensitivity	87.83%
Specificity	97.57%
Accuracy	95.95%

Table 3 Evaluation results

Another evaluation was done to know the misestimation factor. In case of $\overline{E_i}$ (window size 20), if misestimation happened, then the error was just 14 times bigger than one level. This means, in case of misestimation, the level is still reliable to the actual traffic density level.

6. CONLUSION

The evaluation of the algorithm was successful and it fulfilled the expectations. The accuracy was 95.95%. The algorithm performance is lower in case of high traffic density because of physical substances (the energy of the environmental noise become lower when the vehicles slow down during the congestions). This effect and the influence of the sounds, which are similar to the traffic noise, show the limitation of the energy based deterministic approach. Machine learning based approaches using relevant features apart from energy descriptors to reduce the disturbance effects are proposed to increase the robustness of the methods.
The introduced work shows that audio solutions can be used in smart cities and it proves the concept that traffic density estimation based on energy level is a possible and useful application. Currently, the audio sensor is working for more than eight months in Santander, Spain without any problem, which proves the robustness of the device. The advanced processing power, the audio based applications, the developed hardware and software is in line with IoT requirements and contributes to the development of a set of new generation sensors for smart cities.

7. REFERENCES

[1] EU FP7 Project EAR-IT Experimenting Acoustic in Real environment using Innovative Test-beds (ID 318381). Online: http://www.ear-it.eu

[2] EU FP7 Project Smart Santander (ID 257992). Online: http://www.smartsantander.eu.

[3] Gordon, Robert L., Dunn Engineering Associates In association with Siemens Intelligent Transportation Systems, Traffic control systems handbook, Washington, DC : U.S. Dept. of Transportation, Federal Highway Administration, Office of Operations, 2005

[4] Bengang Li, Shu Tao, R.W.Dawson, Evaluation and analysis of traffic noise from the main urban roads in Beijing, Elsevier, Applied Acoustics, 2002, pp. 1137-1142

[5] Enhancing Wireless Sensor Networks with Acoustic Sensing Technology: Use Cases, Applications & Experiments, Danilo Hollosi, György Nagy, Rene Rodigast, Stefan Goetze, Philippe Cousin

[6] Waspmote IoT dataheet. Online: http://www.libelium.com/products/waspmote/hardware/, Last visited: 07.02.2014

[7] Advanticsys IoT Datasheet. Online: http://www.epssilon.cl/files/EPS3000.pdf, Last visited: 07.02.2014

[8] Definitions of the Internet of Things. Online: http://postscapes.com/internet-of-things-definition, Last visited: 07.02.2014

[9] Paulo H. T. Zannin, Fabiano B. Diniz, Wiliam A. Barbosa, Environmental noise pollution in the city of Curitiba, Brazil, Elsevier, Elsevier, Applied Acoustics, 2002, pp. 351-358

[10] Policy Department A: Economic and Scientific Policy, Mapping Smart Cities in the EU, 2014